CHAPTER 6 bis

Distributions

The Dirac "function" has proved extremely useful and convenient to physicists, even though many a mathematician was truly horrified when the Dirac function was described to him: a function that is infinity at zero and zero everywhere else. Eventually it was given a rigorous meaning as a "distribution". Another advantage of distributions is the availability of derivatives and this makes them a useful tool in the theory of partial differential equations. This section introduces the basic concepts and properties.

1. Test functions

Distributions are defined as the linear functionals of a suitable space of *test functions*. The suitable space in view of Fourier theory is the Schwartz space[1] of smooth functions with fast decay. Let us start by recalling the definition of this space, which we equip with a topology.

DEFINITION. *The* **Schwartz space** $\mathcal{S}(\mathbb{R}^d)$ *of test functions is the linear space of all functions* ϕ *in* $C^\infty(\mathbb{R}^d)$ *that satisfy*

$$\sup_{x\in\mathbb{R}^d} |x|^k \Big| \frac{\partial^\alpha}{\partial x^\alpha} \phi(x) \Big| < \infty$$

for any $k \in \mathbb{N}$ *and any multi-index* $\alpha = (\alpha_1, \dots, \alpha_d)$. *The sequence* (ϕ_n) *converges to* ϕ *iff*

$$\sup_{x\in\mathbb{R}^d} |x|^k \Big| \frac{\partial^\alpha}{\partial x^\alpha} \big(\phi_n(x) - \phi(x)\big) \Big| \longrightarrow 0$$

as $n \to \infty$, *for any* $k \in \mathbb{N}$ *and any multi-index* α.

Convergence in the sense of test functions is a very strong property. One can check that the topology induced by the seminorms above turns $\mathcal{S}(\mathbb{R}^d)$ into a Fréchet space, i.e., a complete metrizable locally convex space (a vector space is locally convex when it is equipped with a family a seminorms that separate points). Whoever is not familiar with analytic technicalities can safely ignore this remark; what matters is the definition of convergence of a sequence of test functions.

[1]The French mathematician Laurent Schwartz (1915–2002) created the theory of distributions in 1950–51.

2. Distributions: Definition and examples

Definition. *A* **distribution** *is a continuous linear functional on* $\mathcal{S}(\mathbb{R}^d)$.

These distributions are often called "tempered" and they are not the only ones. The most common space is the dual of the smooth functions with compact support. This space of test functions is smaller, so the dual space is bigger, which is definitely an advantage. On the other hand, the Fourier transform can only be defined for tempered distributions, this is the reason why we restrict our attention to those.

Let $\mathcal{S}'$ denote the space of distributions, i.e., the dual space of $\mathcal{S}(\mathbb{R}^d)$. This is a linear space, and we consider the weak-* topology (pointwise convergence with respect to test functions). That is, the sequence of distributions (T_n) converges to the distribution T whenever

$$T_n(\phi) \longrightarrow T(\phi)$$

for any test function ϕ.

The simplest examples of distributions are locally integrable functions that do not grow too fast: For such a function f, the corresponding distribution T_f is defined by

$$T_f(\phi) = \int_{\mathbb{R}^n} f(x)\,\phi(x)\,\mathrm{d}x.$$

It is clearly linear and continuous. In order for the integrals above to exist for all $\phi \in \mathcal{S}$, we must assume the existence of a number k such that $\int |f(x)|(1+|x|)^{-k}\mathrm{d}x < \infty$. Many regular measures μ on $\mathbb{R}^d$ (with the Borel σ-algebra) give rise to a distribution T_μ, with $T_\mu(\phi) = \int \phi(x)\mathrm{d}\mu(x)$. Again, these integrals exist provided μ does not put too much mass for large $|x|$. Most distributions are not functions, however. Let us review three important examples.

(i) For $x_0 \in \mathbb{R}^d$ and α a multi-index, let

$$T(\phi) = \frac{\partial^\alpha}{\partial x^\alpha}\phi(x_0).$$

The special case $\alpha = 0$ is **Dirac's "function"**,[2] which is really a distribution. Linearity and continuity are easily verified.

(ii) In $d = 1$, Cauchy's **principal value of** $1/x$, denoted $\mathrm{PV}\frac{1}{x}$, is defined by

$$\mathrm{PV}\tfrac{1}{x}(\phi) = \lim_{\varepsilon\searrow 0}\int_{|x|>\varepsilon} \tfrac{1}{x}\phi(x)\,\mathrm{d}x.$$

Notice that $\lim_{\varepsilon\searrow 0}\int_\varepsilon^\infty \frac{1}{x}\phi(x)\,\mathrm{d}x$ is $\pm\infty$ if $\phi(0) \neq 0$. However, this divergence cancels with the divergence of the integral for negative x. This can be seen by writing

$$\int_{|x|>\varepsilon} \tfrac{1}{x}\phi(x)\,\mathrm{d}x = \int_{|x|>1} \tfrac{1}{x}\phi(x)\,\mathrm{d}x + \int_{\varepsilon<|x|<1} \tfrac{1}{x}[\phi(x)-\phi(0)]\mathrm{d}x + \int_{\varepsilon<|x|<1} \tfrac{1}{x}\phi(0)\,\mathrm{d}x.$$

The first integral exists because $\phi(x)$ decays quickly at infinity. The integrand of the second integral is bounded by $\sup|\phi'(x)|$ by the mean-value theorem, so that the integral converges uniformly in ε. Finally, the third integral is zero by

[2]Paul Adrien Maurice Dirac (1902–1984) was a British mathematical physicist who made important contributions in quantum mechanics.

symmetry. This shows that the principal value of $1/x$ is well defined. This is clearly a linear functional. The above decomposition can also be used to prove that $\mathrm{PV}\frac{1}{x}(\phi_n) \to 0$ for any sequence of test functions (ϕ_n) that converges to 0, so this functional is continuous.

(iii) Again in $d = 1$, the distribution $\frac{1}{x+\mathrm{i}0}$ is defined by

$$\tfrac{1}{x+\mathrm{i}0}(\phi) = \lim_{\varepsilon \searrow 0} \int_{-\infty}^{\infty} \frac{\phi(x)}{x + \mathrm{i}\varepsilon} \, \mathrm{d}x;.$$

Convergence of the limit $\varepsilon \to 0$ can be proved in about the same way as for $\mathrm{PV}\frac{1}{x}$.

These three distributions are actually related by the identity (Exercise 5).

$$\tfrac{1}{x+\mathrm{i}0} = \mathrm{PV}\tfrac{1}{x} - \mathrm{i}\pi\delta_0.$$

3. Weak or distributional derivatives

A major advantage of distributions is to generalise the notion of derivatives in a very natural way. Indeed, it is intuitively clear that the derivative of $|x|$ should be $\mathrm{sgn}(x)$, although the usual definition does not apply. Here, the infinite differentiability of test functions is essential.

DEFINITION. *Let $\alpha = (\alpha_1, \dots, \alpha_n)$ a multi-index. The* **distributional derivative** *or* **weak derivative** *of a distribution T is defined by*

$$\big(\partial^\alpha T\big)(\phi) = (-1)^\alpha \, T\Big(\frac{\partial^\alpha}{\partial x^\alpha} \, \phi\Big).$$

Here, $(-1)^\alpha = (-1)^{\alpha_1 + \dots + \alpha_d}$.

In particular, in $d = 1$, $T'(\phi) = -T(\phi')$. Let us see that this definition is indeed an extension of the usual definition of the derivative of a function. If f is a differentiable function and T_f is the corresponding distribution, we have This follows from the definition and an integration by parts:

$$T_f'(\phi) = -T_f(\phi') = -\int_{-\infty}^{\infty} f(x)\phi'(x) \, \mathrm{d}x = \int_{-\infty}^{\infty} f'(x)\phi(x) \, \mathrm{d}x = T_{f'}(\phi).$$

It follows that $T_f' = T_{f'}$. This computation also clarifies the sign that appears in the definition of the distributional derivative.

Let us illustrate this notion in three situations where the original definition of derivative does not apply.

(a) The distributional derivative of $|x|$ is $\mathrm{sgn}(x)$:

$$T_{|x|}'(\phi) = -\int_{-\infty}^{\infty} |x|\phi'(x) \, \mathrm{d}x = -\int_{-\infty}^{0} \phi(x) \, \mathrm{d}x + \int_{0}^{\infty} \phi(x) \, \mathrm{d}x = T_{\mathrm{sgn}(x)}(\phi).$$

(b) The distributional derivative of the Heaviside step function[3] is the Dirac distribution. Let $H(x) = 1$ if $x \geqslant 0$, $H(x) = 0$ if $x < 0$. Then

$$T'_H(\phi) = -T_H(\phi') = -\int_0^\infty \phi'(x)\,\mathrm{d}x = -\phi(x)\Big|_0^\infty = \phi(0).$$

(c) The distributional derivative of the Dirac distribution, noted δ'_x, is easily found:

$$\delta'_x(\phi) = -\delta_x(\phi') = -\phi'(x).$$

Lemma 1. *The map $T \mapsto \partial^\alpha T$ is continuous.*

Proof. If $T_n \to T$, then $\partial^\alpha T_n(\phi) = (-1)^\alpha T_n(\frac{\partial^\alpha}{\partial x^\alpha}\phi) \to (-1)^\alpha T(\frac{\partial^\alpha}{\partial x^\alpha}\phi) = \partial^\alpha T(\phi)$. □

4. Fourier transforms of distributions

We now extend the notion of Fourier transform to distributions. We saw before how to define derivatives of distributions, that coincide with the usual definition when the distribution is a differentiable function. The situation is similar with the Fourier transform. It applies to any distribution, and coincides with the definitions above when the distribution is an L^1 function.

We first need to understand what should be the definition of the Fourier transform of a distribution. If $f \in L^1(\mathbb{R})$, the Fourier transform of T_f should be $\widehat{T_f} = T_{\widehat{f}}$. Consequently,

$$\begin{aligned}\widehat{T_f}(\phi) = T_{\widehat{f}}(\phi) &= \int_{-\infty}^\infty \widehat{f}(k)\phi(k)\,\mathrm{d}k = \int_{-\infty}^\infty \Big[\int_{-\infty}^\infty \mathrm{e}^{-2\pi\mathrm{i}kx}\, f(x)\,\mathrm{d}x\Big]\phi(k)\,\mathrm{d}k \\ &= \int_{-\infty}^\infty \Big[\int_{-\infty}^\infty \mathrm{e}^{-2\pi\mathrm{i}kx}\,\phi(k)\,\mathrm{d}k\Big] f(x)\,\mathrm{d}k = \int_{-\infty}^\infty f(x)\widehat{\phi}(x)\,\mathrm{d}x = T_f(\widehat{\phi}).\end{aligned}$$

The interchange of integrals is justified by Fubini theorem (check it!). This shows that for any distribution that is given by an L^1 function, we should have

$$\widehat{T_f}(\phi) = T_f(\widehat{\phi}).$$

The idea for general distributions is to use this property as a definition. But this works only if the Fourier transform of a test function is a test function, and this is the main reason behind using Schwartz functions as test functions.

Definition. *The* **Fourier transform** *of the distribution $T \in \mathcal{S}'$ is the distribution $\widehat{T} \in \mathcal{S}'$ that satisfies*

$$\widehat{T}(\phi) = T(\widehat{\phi})$$

for any $\phi \in \mathcal{S}(\mathbb{R}^d)$.

---

[3]Oliver Heaviside (1850–1925), a British who had strong opinion about the Euclidean geometry taught in school: "It is shocking that young people should be addling their brains over mere logical subtleties, trying to understand the proof of one obvious fact in terms of something equally... obvious." A great applied mathematician nonetheless!

The functional $\widehat{T}$ defined above is obviously linear, since T and the Fourier transform are linear. Is it continuous? If $\phi_n \to \phi$ in the sense of test functions, we can check that $\widehat{\phi}_n \to \widehat{\phi}$ (also in the sense of test functions), see Exercise 10, and therefore

$$\widehat{T}(\phi_n) = T(\widehat{\phi}_n) \longrightarrow T(\widehat{\phi}) = \widehat{T}(\phi).$$

Then $\widehat{T}$ is continuous indeed.

The definition above allows to take the Fourier transform of any distribution. In particular, any locally integrable function (which does not diverge too quickly at infinity) has now a Fourier transform, not only L^1 and L^2 functions. It may be, however, that the Fourier transform of a non-L^1 or non-L^2 function is a distribution that is not a function.

Lemma 2. *The map $T \mapsto \widehat{T}$ is continuous.*

Proof. If $T_n \to T$, then $\widehat{T}_n(\phi) = T_n(\widehat{\phi}) \to T(\widehat{\phi}) = \widehat{T}(\phi)$. $\square$

This lemma is useful; it sometimes easier to compute the Fourier transform of an approximating sequence and to take the limit. In Exercise 3, this is done with the function $\frac{1}{|x|}$ in dimension $d = 3$.

It turns out that the Fourier transform is a bijection on the space of distributions. In order to define the inverse transform, let

$$\phi^\vee(k) = \int_{\mathbb{R}^d} \mathrm{e}^{2\pi \mathrm{i} kx}\, \phi(x)\, \mathrm{d}x.$$

Proposition 3.

(a) *The map $T \mapsto \widehat{T}$ is a bijection $\mathcal{S}' \to \mathcal{S}'$.*
(b) *The inverse map is $T \mapsto T^\vee$, where $T^\vee(\phi) = T(\phi^\vee)$.*

Proof. The maps $\widehat{\cdot}$ and $\cdot^\vee$ are both defined on the whole of $\mathcal{S}'$. We have

$$\widehat{T}^\vee(\phi) = \widehat{T}(\phi^\vee) = T(\widehat{\phi^\vee}) = T(\phi).$$

Then $T = \widehat{T}^\vee$, and a similar argument leads to $T = \widehat{T^\vee}$. This implies that the maps $\widehat{\cdot}$ and $\cdot^\vee$ are both onto. If $T(\phi) = 0$ for all ϕ, then $\widehat{T} = 0$ and $\widehat{T}^\vee = 0$, so $T = 0$ and the map $\widehat{\cdot}$ is one-to-one. $\square$

We have already seen the effect of the Fourier transform on functions of the type xf and f'. The distributional Fourier transform is similar. We denote $x^\alpha T$ (or $k^\alpha T$) the distribution such that $(x^\alpha T)(\phi) = T(x^\alpha \phi)$. It is well defined since $x^\alpha \phi \in \mathcal{S}(\mathbb{R}^d)$.

Proposition 4.

(a) $\partial^\alpha \widehat{T} = (-2\pi \mathrm{i})^\alpha\, \widehat{x^\alpha T}$.
(b) $\widehat{\partial^\alpha T} = (2\pi \mathrm{i}\, k)^\alpha\, \widehat{T}$.

PROOF. For item (a),

$$(\partial^\alpha \widehat{T})(\phi) = (-1)^\alpha \widehat{T}(\partial^\alpha \phi) = (-1)^\alpha T(\widehat{\partial^\alpha \phi}) = (-1)^\alpha T((2\pi\mathrm{i}\, k)^\alpha \widehat{\phi})$$
$$= (-2\pi\mathrm{i})^\alpha (x^\alpha T)(\widehat{\phi}) = (-2\pi\mathrm{i})^\alpha \widehat{x^\alpha T}(\phi).$$

Item (b) is similar:

$$\widehat{(\partial^\alpha T)}(\phi) = \partial^\alpha T(\widehat{\phi}) = (-1)^\alpha T(\partial^\alpha \widehat{\phi}) = (-1)^\alpha T((-2\pi\mathrm{i})^\alpha \widehat{x^\alpha \phi})$$
$$= (2\pi\mathrm{i})^\alpha\, \widehat{T}(x^\alpha \phi) = (2\pi\mathrm{i}\, k)^\alpha\, \widehat{T}(\phi).$$

□

5. The Poisson equation

The Poisson equation[4] is a differential equation in three dimensions for the gravitational potential or the Coulomb potential. The equation is

$$-\Delta u = f,$$

where Δ denotes the Laplacian,

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}.$$

The function u is the Coulomb potential if f is the distribution of charge; u is the negative of the gravitational potential if f is the distribution of mass.

We suppose here that $f \in \mathcal{S}(\mathbb{R}^3)$, which simplifies the study. We first search for a solution using formal calculations. The Fourier transform of the equation is

$$4\pi^2 |k|^2 \widehat{u}(k) = \widehat{f}(k).$$

We find $\widehat{u}(k) = \frac{1}{4\pi^2|k|^2}\widehat{f}(k)$. The inverse Fourier transform of $\frac{1}{4\pi^2|k|^2}$ is $\frac{1}{4\pi|x|}$ (see Exercise 3), which suggests that

$$u(x) = \Big(\frac{1}{4\pi|x|} * f\Big)(x) = \int \frac{1}{4\pi|x-y|} f(y)\mathrm{d}y = \int \frac{1}{4\pi|y|} f(x-y)\mathrm{d}y. \qquad (\Diamond)$$

For $f \in \mathcal{S}(\mathbb{R}^3)$, these integrals are well-defined for all x and u is C^∞. The goal is to show that for any Schwartz function f,

$$f(x) = -\Delta u(x) = \int \big(-\Delta \tfrac{1}{4\pi|x-y|}\big) f(y)\mathrm{d}y = \int \big(-\Delta \tfrac{1}{4\pi|y|}\big) f(x-y)\mathrm{d}y.$$

This suggests that the Laplacian of $1/|x|$ is proportional to a Dirac distribution. We now establish this rigorously.

LEMMA 5. *Let T be the distribution associated with $\frac{1}{4\pi|x|}$. Then*

$$-\Delta T = \delta_0.$$

[4]Siméon-Denis Poisson (1781–1840) was a French whose innovative works in applied mathematics and in mathematical physics include potential theory in electromagnetism, for which the "Poisson equation" plays an important rôle. According to Poisson lLife is good for only two things, discovering mathematics and teaching mathematics!

The function $\frac{1}{4\pi|x|}$ is called the "Green function"[5] of the Poisson equation.

Proof. By Proposition 3, two distributions are equal iff their Fourier transforms are equal. We have

$$\widehat{-\Delta T} = 4\pi^2|k|^2\widehat{T} = 4\pi^2|k|^2 T_{\frac{1}{4\pi^2|k|^2}} = T_1 = \widehat{\delta_0}.$$

$\square$

Given $\phi \in \mathcal{S}(\mathbb{R}^3)$, let ϕ_x denote the function $\phi_x(y) = \phi(x-y)$. Clearly, the map $\phi \mapsto \phi_x$ is a bijection $\mathcal{S} \to \mathcal{S}$. The function u in ($\diamondsuit$) can be written as

$$u(x) = T(f_x).$$

Furthermore, we have

$$-\Delta u(x) = -\int \frac{1}{4\pi|y|}(\Delta f)(x-y)\mathrm{d}y = -T(\Delta f_x).$$

We used the relation $\Delta f(x-y) = \Delta f_x(y)$. By the definition of weak derivatives and Lemma 5, we have

$$T(\Delta f_x) = \Delta T(f_x) = \delta_0(f_x) = f_x(0) = f(x).$$

We have just proved that $-\Delta u(x) = f(x)$.

Exercise 1. Let $d = 1$. Show that the Fourier transform of $\frac{1}{x^2+\mu^2}$ is $\frac{\pi}{\mu}\,\mathrm{e}^{-2\pi\mu|k|}$ (use contour methods).

Exercise 2. Let $d = 1$. Show that the Fourier transform of $\mathrm{e}^{-2\pi\mu|x|}$ is $\frac{\mu}{\pi}\,\frac{1}{k^2+\mu^2}$.

Exercise 3. Let $d = 3$.
(a) Show that the Fourier transform of $\frac{1}{|x|}\,\mathrm{e}^{-2\pi\mu|x|}$ is $\frac{1}{\pi}\,\frac{1}{k^2+\mu^2}$.
(b) Find the Fourier transform of the function $1/|x|$. Since this function is not in any L^p space, explain the meaning of the Fourier transform.

Exercise 4. In $d = 1$, describe the distribution $x^k\partial^\ell\delta_0$ for every integers $k, \ell \geqslant 0$.

Exercise 5. Prove that $\frac{1}{x+\mathrm{i}0} = \mathrm{PV}\frac{1}{x} - \mathrm{i}\pi\delta_0$. Hint: use $\frac{1}{x+\mathrm{i}\varepsilon} = \frac{x}{x^2+\varepsilon^2} - \frac{\mathrm{i}\varepsilon}{x^2+\varepsilon^2}$.

Exercise 6. "When differentiating a function with a jump, one picks up a delta multiplied by the height of the jump." Rewrite this sentence in mathematical language. Prove it!

Exercise 7. Compute rigorously the Fourier transforms of
(a) δ_{x_0}.
(b) x.
(c) $x\delta_0$.

[5]The English George Green (1793–1841) was too modest and even he did not understand the importance of his 1828 essay on the mathematical analysis of electricity and magnetism. Much later, this essay would be read by Lord Kelvin, who shared his excitement with Liouville and Sturm.

Exercise 8. Compute rigorously the Fourier transform of $\frac{1}{x+\mathrm{i}0}$. Hint: Obtain the following expression:

$$\widehat{\tfrac{1}{x+\mathrm{i}0}}(\phi) = \lim_{\varepsilon \searrow 0} \lim_{R\to\infty} \int_{-R}^{R} \tfrac{1}{x+\mathrm{i}\varepsilon} \left[\int_{-\infty}^{\infty} \mathrm{e}^{-2\pi \mathrm{i} kx}\, \phi(k)\, \mathrm{d}k \right] \mathrm{d}x.$$

Then use Fubini and contour methods.

Exercise 9. Find the Fourier transform of $\mathrm{PV}\frac{1}{x}$. You may use the results of Exercises 5 and 8 if useful.

Exercise 10. Show that $\phi_n \to \phi$ implies $\widehat{\phi}_n \to \widehat{\phi}$, where convergence is in the sense of the Schwartz space.

References and further reading

Gerald B. Folland, *Real Analysis*, Wiley, 1999.
J. K. Hunter and Bruno Nachtergaele, *Applied Analysis*, World Scientific, 2001.
Elliott H. Lieb and Michael Loss, *Analysis*, AMS, 2001.
Michael Reed and Barry Simon, *Functional Analysis*, Elsevier, 1980
MacTutor History of Mathematics, http://www-history.mcs.st-andrews.ac.uk/history/